DOCUMENT RESUME

ED 218 302 TM 820 332

AUTHOR Stodolsky, Susan S.
TITLE Combining Evidence Between and Within Educational Research Studies.
INSTITUTION Chicago Univ., Ill. Dept. of Education.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Mar 82
NOTE 11p.; Paper presented at the Annual Meeting of the American Educational Research Association (66th, New York, NY, March 19-23, 1982).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Classroom Observation Techniques; *Classroom Research; Data Analysis; Data Collection; *Educational Research; Elementary Secondary Education; Research Methodology; *Research Problems

ABSTRACT

Problems associated with combining observations taken from intact classrooms were examined. Clarification of basic entities to be studied was a critical step in research planning. Classroom observers usually study instructional arrangements and other classroom phenomena. The author used activity structure and its component activity segments as the focus for study. An activity segment was defined as a unique time block in a lesson which occured in a fixed physical setting. Segments have instructional forms designed for the accomplishment of specific tasks. Daily activities in a classroom conform with the general type of curricular approach and goals. Activity structures are produced to accomplish certain goals. When observational data is characterized by distributions which are bimodal or have many zeroes or non-occurences, summary with means and standard deviations may result in erroneous conclusions. The author observed 20 fifth-grade math classes from diverse school districts. He determined that combining data within two clusters, highly individualized and whole class programs, provided more meaningful and interpretable results. Researchers and writers need to include background descriptive information about the programs and contexts which they have studied to facilitate sorting of data. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

S. S. Stodolsky

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED218302

TM 820 332

Combining Evidence Between and Within Educational Research Studies[1,2]

Susan S. Stodolsky, University of Chicago

The use of direct observation for studying classroom process and learning outcomes has increased noticeably in the last decade. Studies of schooling as it actually occurs in classrooms are being conducted to fill our need for timely descriptive information and to help operationally define actual educational treatments. However, observational studies of classroom processes require new methodological skills and strategies.

One particular aspect of observational methodology will be my focus in this paper. I will address problems associated with combining observations taken from intact classrooms. In considering this topic I will draw parallels between combining data within a given observational study and that of combining data from a number of different studies as is done in quantitative research synthesis.

Direct observation of classrooms is executed with a wide variety of specific observational methods. For example, teacher-student interactions have often been observed using such instruments as the Flanders Interaction Analysis Category system (Flanders, 1960) or the Brophy and Good system (1970). Another approach is the use of narrative recording or specimen records as in studies of third-grade classrooms by Gump (1967), videotaping of whole lessons analyzed for critical incidents and judgements of teacher characteristics (Kounin, 1970), and other adaptations of open recording methods such as I have used in research on fifth-grade math and social studies classes (Stodolsky, 1981).

---

[1]Paper presented at the annual meeting of the American Educational Research Association, New York, March, 1982.

[2]Research support was provided by the National Institute of Education. Publication of the Department of Education, Finance and Productivity Center, University of Chicago.

A variety of decisions must be made when conducting non-experimental field studies of classroom processes. Obviously, the study must be framed and operationalized with clear questions and hypotheses in mind and an understanding of the kinds of generalizations which are desired. A critical step in the research planning process is clarification of the nature of the phenomena to be studied. As Dunkin and Biddle (1974) have pointed out, even the most obvious sounding terms can be extremely ambiguous when operationalization is required. For example, the terms "lesson," "interaction" and even "student" may prove difficult to define in certain contexts. Curricular labels such as open education or individualized instruction often connote various and sometimes conflicting meanings. Clarification of the basic entities to be studied deserves careful attention because findings are more easily interpreted if the phenomena being studied are clearly defined.

Once such clarification of questions and phenomena is achieved, observational researchers must decide on some sampling plan. Decisions are made about what to observe, how to observe, when to observe, and how long to observe. These decisions are made with an eye toward the desired generalizations but also within practical constraints such as available resources and teachers' willingness to admit outsiders to their classrooms.

The clarification process involved in planning observational research is somewhat parallel to initial work which must be done in effectively planning a quantitative research synthesis. The parallel is strongest in connection with the process of deciding what studies to include in a synthesis. This decision must be made by considering the construct equivalence of both treatments and measures in diverse studies. I will return to these parallels later in this paper.

Once conducted, what kinds of results do observational studies produce? Typical are findings which describe the frequencies and/or rates of occurence of certain behaviors and other learning conditions. Flander's "law of two-thirds" which asserts that 2/3 of the time in classrooms is spent in talk and 2/3 of the talk is done by the teacher is an example. Another example, from a recent British study of junior schools (Galton, Simon, and Croll, 1981) is that 10 percent of all activities observed across classes in a year's time are cooperative group work. In my research (Stodolsky, 1981) in fifth-grade math and social studies classes we found that group work in social studies occupies 10.7 per cent of student time across 19 different classes observed for two weeks each. We also found that individualized seat work in mathematics occupies 13.7 per cent of student time across 20 different math classes we observed for two weeks each.

These statements about classroom phenomena and their occurence are meant to inform us in some way about the typical experiences of teachers and students as well as the conditions for teaching and learning which are present in our schools. Often an additional step is taken in which such variables are correlated with learning outcomes in order to construct a chain of connections between educational processes and educational effects.

The issue to be raised is the appropriateness of aggregating data taken from a variety of intact classes, lessons, or educational environments. What substantive and quantitative assumptions should be met in order to summarize data across classes using standard descriptive statistics such as means and standard deviations? How do we decide if the interpretation of such data will be valid? For example, shall we assume that children in British junior schools do group work activities 10% of the time?

I address this query first from a theoretical and conceptual perspective about the nature of instruction. The variables usually studied by observers are aspects of instructional arrangements and other classroom phenomena. Instructional arrangements are significantly and fundamentally constrained by educational purpose and organization and are highly interdependent phenomena. In my own work I view educational settings through the use of an ecological perspective, using the activity structure and its component activity segments as the focus for study. The concept of an activity segment is derived from the idea of a behavior setting, a fundamental concept in ecological psychology (Barker, 1968). An activity segment is part of a classroom activity structure which has a particular instructional format, participants, materials, behavioral expectations and goals, and space-time boundaries. A segment is defined as a unique time block in a lesson and occurs in a fixed physical setting. Segments can occur singly as when the whole class is involved in the same activity such as a teacher-led recitation or simultaneously as when four groups of children are working on tasks.

Knowing a particular activity segment exists does not permit one to predict all the molecular behavioral transactions in the classroom, but it does significantly limit the likelihood of seeing certain behaviors as opposed to others. Segments have instructional forms which are suited to the accomplishment of certain tasks. While the detailed connections between form and tasks must be investigated in more depth, it is clear that a series of contextual and pedagogical factors affect the configuration of behaviors and activities which occur in a classroom at any given time. For example, knowing that children are participating in a whole-class recitation will suggest a context in which teacher talk is likely

to predominate and child-child interaction is likely to be very infrequent. On the other hand knowing that children are doing seatwork in a math class will mean that teacher-child interactions are likely to be private and that children will be writing as a primary activity. In such cases the cognitive level of the activity is not fixed, although prior experience and research would indicate that certain levels of question-asking are more likely to occur in recitations and certain types of problems are likely to be solved in a seatwork setting.

Beyond knowing the properties of a particular setting and the constraints that it imposes, the fact that settings and segments are designed to accomplish certain curricular purposes must also be considered. The daily activities in a classroom are ordered and cohere at many levels, one important one is in conformity with the general type of curricular approach and goals. Activity structures are produced to accomplish certain goals. For example, a teacher attempting to implement an open education program is likely to structure the day to enable children to make a variety of choices about the activities they will pursue. An individualized math program will probably consist of seatwork sessions in which children work at their own pace, possibly interspersed with tests and teacher conferences. In social studies, certain curricula are predicated on the use of peer work groups. For example, the MACOS curriculum contains many activities which are to be carried out by a group of children working cooperatively.

With programatic constraints operating, it is clear that classrooms must look quite different from one another. In fact this is the essence of educational diversity. Therefore it is insufficient to think of the variability as connected to or associated with some typical average. In fact it may often be necessary to recognize interrelated patterns which must be distinguished one from the other.

Our own data provide some convenient illustrations of this point. I indicated earlier that we observed twenty fifth-grade math classes from diverse school districts for two consecutive weeks. We found that individualized seatwork in which children work at their own pace and use testing and diagnostic devices accounted for 13.7 percent of student occupancy time across all classes and occasions. Thus one might expect that walking into a math class at the fifth grade, the observer would see children occupied with individualized seatwork on one of every eight occasions. However, an inspection of the data would show that this summary statistic is a very poor summary indeed. Actually 11 of 20 classes never used the individualized seatwork approach. Four classes had this arrangement less than five percent of the time (basically on one occasion) while in the other five classes it occupied 71%, 70%, 55%, 32% and 19% of student time. Essentially we have a three-way clustering on this variable: 1) The program (individualized seatwork) is not used at all. 2) The program is used infrequently for a special purpose. 3) The program is the major curricular approach utilized. To adopt another type of language, these clusters would seem to reflect at least three different curricular treatments. To the extent that these classes are implementing different curricular approaches and treatments, it would seem essential to both substantively and statistically seperate them for purposes of analysis. It does not seem appropriate, at this level of analysis, to think of them as forming some sort of continuum to be used for descriptive or predictive purposes. They are coherent educational entities which are operating with different assumptions, practices, and possibly different goals.

A similar configuration can be found with regard to the occurence of peer instructional work groups in social studies. While we have an

average occupancy time of 10.7 percent in group work settings across 19 classes, 9 classes are never observed using groupwork settings, 3 classes use them occasionally, and seven use them a substantial amount suggesting that they are a dominant instructional form in regard to the curriculum being utilized. Similarly Galton, Simon and Croll (1981) recently found that groupwork in British junior schools did not occur at all for approximately 90 percent of the children, even though groupwork accounted for 10 percent of all activities observed across classes.

Observational data is often characterized by distributions which are either bimodal or which have many zeros or non-occurences. To summarize such data with means and standard deviations or to use such variables as predictors is perhaps to assume too much (or the wrong things) about our knowledge of the phenomena. In particular, when variables are distributed in ways we have just described for groupwork or individualized seatwork, displaying average occurences or correlating their occurence with outcome variables suggest that we have a variable whose sheer frequency will show an orderly relation to other variables. But in these cases and many other observational data contexts low or no frequency of occurence on a given variable means that an <u>entirely different educational program</u> is in operation which must be evaluated as such.

It seems much more important and appropriate to attempt to identify treatments which are in fact similar on substantive grounds in order to better combine data. Thus it would seem more appropriate to talk about our mathematics data by first indicating that there are two major methods we have observed in mathematics instruction. Programs which are highly individualized and operate on a self-paced plan with uniform goals for all children, and programs which are enacted for the whole class assuming

that all children will keep pace. In some of the latter classes a small number of advanced or delayed students may be seperated in order to go beyond the whole class goals or for remediation and this is often done with individualized approaches.

Combining data within these two clusters of classrooms would seem both more meaningful and interpretable in terms of either description or prediction. Variations within these clusters would allow one to investigate the relative efficacy of certain transactions or arrangements, but within the framework of a pedagogical context that was really operating in classrooms. While curricular contrasts may be helpful and appropriate on occasion, studies often fail because so many treatment configurations are combined that patterns of relations cannot emerge.

It is in this respect that the problems faced by persons conducting quantitative research syntheses are very similar to those of the observational researchers. How does one decide when it is appropriate to combine data from a variety of studies? What does it mean to be measuring the same effect across studies? Light and Pillemer (1982) have suggested the necessity for including both narrative and statistical indicators of treatment similarities in order to interpret quantitative research syntheses. Giaconia and Hedges (1981) have illustrated analyses which consider curricular factors in combining studies of open education. They identified seven features identified with the practice of open education, four of which were deemed essential for a well-implemented open education program. In conducting a quantitative research synthesis every study was coded for the seven features they identified. In doing this they operationally defined the treatments that were represented in the studies they were reviewing. By examining features of open education, they discovered that

certain effects were strongly associated with the occurence of the program features. In particular, all studies which showed large effect sizes for self concept as an educational outcome were found to have the four essential features they identified. The Giaconia and Hedges (1981) example illustrates the actual steps one might take to assure treatment similarity in studies before looking for consistent effects across the studies. This step seems as necessary within observational research studies as it does in combining studies.

In both observational studies and quantitative research synthesis, more detective work must be done about the nature of the educational phenomena being studied. Reserachers and writers would advance the state of the art by including more in the way of background descriptive information about the programs and contexts which they have studied to facilitate sorting and resorting of data. A recognition that educational settings really are multivariate in the most fundamental sense seems essential to advance both descriptive work in education and subsequent analyses and reanalyses of existing data.

References

Barker, R. G. Ecological Psychology: Concepts and Methods for Studying the Environment of Human Behavior. Stanford: Stanford University Press, 1968.

Brophy, J. E. & Good, T. L. Teachers communication of differential expectations for children's classroom performance: Some behavioral data. Journal of Educational Psychology, 61 (5), 1970, 365-374.

Dunkin, M. J. & Biddle, B. J. The Study of Teaching. New York: Holt, Rinehart, and Winston, 1974.

Flanders, N. A. Interaction Analysis in the Classroom: A Manual for Observers. Ann Arbor: University of Michigan, 1960.

Galton, M., Simon, B. & Croll, P. Inside the Primary Classroom. London: Routledge & Kegan Paul, 1981.

Gump, P. V. The Classroom Behavior Setting: Its Nature and Relation to Student Behavior. Final Report, Project No. 2453 U. S. Office of Education. Lawrence, Kansas: University of Kansas Press, 1967.

Giaconia, R. M. & Hedges, L. V. Identifying features of effective open education. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April, 1981.

Kounin, J. S. Discipline and Group Management in Schools. New York: Holt, 1970.

Light, R. J. & Pillemer. Numbers and narrative: Combining their strength in research reviews. Harvard Educational Review, 52 (1), 1982, 1-26.

Stodolsky, S. S. Subject matter constraints on the ecology of classroom instruction. Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April, 1981.